

VGG16 모델을 기반으로 심음의 Mel Spectrogram을 사용한 심장 질환 분류

Classification of Heart disease using Mel-Spectrogram of PCG based on VGG16 model

저자 (Authors)	장윤영, 배진희, 임준식 Yun Young Jang, Jin Hee Bae, Jun-S. Lim
출처 (Source)	한국정보과학회 학술발표논문집 , 2019.12, 1537-1539(3 pages)
발행처 (Publisher)	한국정보과학회 The Korean Institute of Information Scientists and Engineers
URL	http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE09301992
APA Style	장윤영, 배진희, 임준식 (2019). VGG16 모델을 기반으로 심음의 Mel Spectrogram을 사용한 심장 질환 분류. 한국정보과학회 학술발표논문집, 1537-1539
이용정보 (Accessed)	상명대학교 천안캠퍼스 203.237.183.*** 2021/10/25 18:51 (KST)

저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

VGG16 모델을 기반으로 심음의 Mel Spectrogram을

사용한 심장 질환 분류

장윤영^o, 배진희, 임준식
가천대학교 컴퓨터공학부

jetevoisu@gmail.com, rkcsjwsl@daum.net, jslim@gachon.ac.kr

Classification of Heart disease using Mel-Spectrogram of PCG

based on VGG16 model

Yun Young Jang^o, Jin Hee Bae, Jun-S. Lim
Dept. of Computer Engineering, Gachon University

요 약

심음(PCG)은 심장질환을 진단하는 것에 있어서 중요한 실마리를 제공한다. 하지만 심장질환이 없는 사람도 심잡음은 존재하기 때문에 심음을 이용하여 심장질환의 유무를 파악하는 것은 매우 어렵다. 본 논문은 PhysioNet/CinC Challenge 2016에서 제공한 심음 데이터를 사용하였다. 심음 데이터의 멜 스펙트로그램(Mel-Spectrogram) 이미지를 기반으로 정상/비정상으로 분류한다. 심음을 5초 단위로 분리하여, 분리된 5초간의 심음 데이터를 Mel-Spectrogram으로 변환한다. 각 심음으로부터 추출한 Mel-Spectrogram을 입력값으로 하여 VGG16 모델을 학습시킨다. 5-fold 교차검증 방법으로 모델을 평가한 결과, Accuracy는 0.93, Sensitivity는 0.8598, Specificity는 0.9527 그리고 Precision은 0.8545의 결과를 보였다.

1. 서 론

심음은 심장질환을 예측하는 것에 있어서 중요한 정보를 제공한다. 하지만 심장질환이 없는 정상인의 심장에서도 정상적이지 않은 소리가 흔히 들리기도 하는데 이를 심잡음이라고 한다. 이러한 심잡음은 심음으로 심장질환의 유무를 파악하는 것에 있어서 큰 어려움을 준다. 따라서 심음 소리로부터 특정 소리가 심잡음인지 심장질환에 의한 잡음인지 구분하는 것은 매우 중요하다[1].

소리는 일종의 시계열적 특징을 지닌 데이터로써, 이러한 상태의 소리에서는 정상과 비정상을 구분하기 위한 특징들을 찾아내기 쉽지 않다. 하지만 Mel-Spectrogram은 시간에 따른 주파수의 변화와 진폭의 변화를 색상으로 보여주기 때문에 시간에 따른 음성의 변화를 시각적으로 볼 수 있다. [2]

본 연구에서는 심음의 Mel-Spectrogram을 효과적으로 분류하기 위하여 VGG16 모델을 사용한다. VGG16 모델은 옥스퍼드 대학의 연구팀이 개발한 CNN의 기반의 모델로써 이미지 인식 대회에서 준우승을 차지할 만큼 이미지 분류에 대해 좋은 성능을 보여준다.[3] 하지만 사전에 훈련이 되어있는 이 모델은 훈련에 사용된 데이터

셋이 심음의 Mel-Spectrogram과 맞지 않기 때문에 우리는 이 모델에 대해 일부 가중치 값을 초기화시킨 뒤 다시 학습시키는 방법을 사용한다. 이러한 방법을 통해 VGG16 모델이 가지고 있는 이미지 분류의 성능을 저하시키지 않고 모델을 특정 데이터 특성에 맞게 변형할 수 있다.

따라서 본 논문에서는 VGG16 모델을 이용해 심음을 정상/비정상 두 가지 범주로 분류하여 심장질환을 가진 환자를 파악할 수 있는 알고리즘을 제시한다. 예측 결과는 PhysioNet/CinC Challenge 2016에서 CNN과 Spectrogram을 사용한 모델[4]과 비교하여 분석한다.

2. Challenge Data

PhysioNet은 10년 이상의 기간에 걸쳐 7개국의 연구팀들이 수집한 심음 데이터베이스를 제공했다. 챌린지의 훈련 세트(a ~ f)는 764명의 건강한 피험자와 환자로부터 5에서 120초 이상 지속되는 총 3,153개의 심음을 녹음한 것이며, 테스트 세트(b ~ e, + g 및 i)는 308명의 건강한 피험자와 환자로부터 5에서 120초 이상 지속되는 총 1,277개의 심음을 녹음했다. 심음 데이터는 전문

가에 의해 정상/비정상으로 분류되었다.[5] 본 논문에서는 훈련 세트(a~f)를 데이터로 사용한다.

3. 음성 데이터 특징 추출

3.1 음성 데이터 분할

PhysioNet에서 제공한 심음 데이터는 그 길이가 5초에서 120초 사이로 매우 다양하며 데이터의 양이 적다. 따라서 우리는 음성 데이터를 분할함으로써 데이터의 길이를 정규화시키고 학습데이터의 양을 증가시켰다. 음성 데이터를 나누는 window 크기의 기준은 모든 음성 데이터의 길이 중 가장 짧은 5초를 기준으로 하였다. 모든 음성 데이터를 5초 단위로 중복되는 영역 없이 분할하여 3158개의 비정상 음성 데이터와 9857개의 정상 음성 데이터를 얻었다.

3.2 Mel-Spectrogram으로 변환

1차원의 심음을 2차원의 시간에 따른 주파수의 모습인 Mel-Spectrogram으로 변환해주기 위해 STFT(Short Time Fourier Transform)와 Mel Filterbank를 이용한다. 분할한 심음 데이터에 대해서 25ms의 윈도우(window) 크기와 10ms의 간격을 기준으로 STFT를 적용한다. STFT를 적용하여 얻어진 값에 Mel Filter Bank를 적용하고 각 필터의 에너지 값을 합한다. 이를 이미지로 나타낸 것이 Mel-Spectrogram이다.[2][6] 이러한 과정을 거쳐 비정상 Mel-Spectrogram 3158개와 정상 Mel-Spectrogram 9857개를 추출했다.



그림 1 5초 간격의 정상/비정상 Mel-Spectrogram

4. VGG16 네트워크

VGG는 ImageNet 데이터 셋을 사전 학습한 Convolutional neural network이며 사전 훈련된 합성곱(Convolution), 풀링(Pooling)층과 새로운 완전연결층(fully connected Layer)을 이용하여 새로운 데이터에 대한 특징을 추출한다. 네트워크의 성능을 높이기 위해서는 미세조정(fine tuning)을 해야 한다. 새로운 문제에 재사용하도록 수정

이 필요한 것은 구체적 특성이므로 상위층인 Conv_Block5 층을 미세조정(fine tuning) 했다.[7]

그림 2와 같이 네트워크는 Convolution Layer과 Pooling Layer(Max Pooling)으로 이루어져 있으며 16개의 층으로 이루어진 VGG16모델이다. 입력 데이터는 100x100 크기의 Mel-Spectrogram을 사용했고 정상, 비정상 범주로 나누는 이진분류 모델이다. 데이터와 모델의 검증은 CNN과 Spectrogram을 사용한 모델[4]과 동일하게 5-fold 교차 검증(5-fold cross validation)을 사용하여 진행하였다. 모델 구현은 오픈소스 신경망 라이브러리인 케라스(Keras)를 사용했으며 옵티마이저는 RMSprop을 이용하여 가중치를 업데이트 했다.

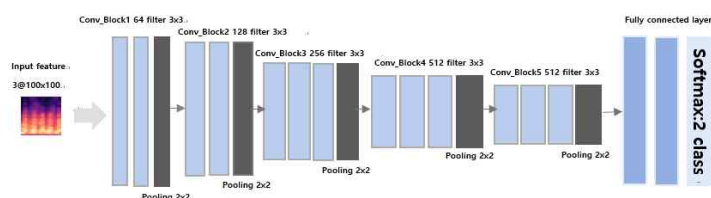


그림 2 VGG16 네트워크 구조

5. 결과

Mel-Spectrogram의 이미지를 VGG16 모델로 5-fold 교차검증을 진행한 결과, 정확도는 0.98, 민감도는 0.8598, 특이도는 0.9527 그리고 정밀도는 0.8545였다. CNN 모델과 Spectrogram을 이용한 실험 결과와 비교하였을 때, 민감도는 0.04 높게 나왔지만 특이도는 0.01정도 낮은 수치를 보여주었다. 하지만 PhysioNet/CinC Challenge 2016에서 성능을 평가하는 기준인 Score 점수를 비교하였을 때에는 우리가 제시하는 모델의 Score 점수는 0.9로 CNN을 사용한 실험보다 더 좋은 성능을 보여준다고 판단할 수 있다.

VGG16	
Sensitivity	0.85
Specificity	0.95
Score	0.9
CNN을 사용한 모델[4]	
Sensitivity	0.81
Specificity	0.96
Score	0.885

표 1 VGG16과 CNN을 사용한 결과비교

6. 결론

Blog , 2016

우리는 PhysioNet/CinC Challenge 2016에서 제공하는 데이터를 이용하여 심음을 정상/비정상으로 분류하는 실험을 진행하였다. 심음을 Mel-Spectrogram으로 이미지화 한 후 VGG16 모델을 사용하여 검증한 결과, Accuracy는 0.93, Sensitivity와 Specificity는 각각 0.8598, 0.9527을 보여줌으로써 CNN과 Spectrogram을 사용한 결과[4]보다 더 높은 성능을 보이는 것을 확인할 수 있었다.

참고문헌

- [1] Kwak Chul, Kwon Oh Wook, “Heart Sound-Based Cardiac Disorder Classifiers Using an SVM to Combine HMM and Murmur Scores” in *The Journal of the Acoustical Society of Korea*, vol. 30, issue. 3, 149–157, 2011
- [2] Koustav Chakraborty, Asmita Talele, Prof. Savitha Upadhyaya, “Voice Recognition Using MFCC Algorithm” in *International Journal of Innovative Research in Advanced Engineering (IJIRAE)*, vol.1, Issue 10, 2014
- [3] Karen Simonyan, Andrew Zisserman, “Very deep convolutional networks for large-scale image recognition” in *arXiv preprint arXiv*, 2014
- [4] Tanachat Nilanon , Jiayu Yao , Junheng Hao , Sanjay Purushotham , Yan Liu, ” Normal / Abnormal Heart Sound Recordings Classification Using Convolutional Neural Network”, in *2016 Computing in Cardiology Conference (CinC)* , 585–588, 2016
- [5] Gari D Clifford¹, Chengyu Liu, Benjamin Moody, David Springer, Ikaro Silva, Qiao Li, and Roger G. Mark, “Classification of Normal/Abnormal Heart Sound Recordings: the PhysioNet/Computing in Cardiology Challenge 2016” in *2016 Computing in Cardiology Conference (CinC)*, 609–612, 2016
- [6] Jonathan Rubin, Rui Abreu, Anurag Ganguli, Saigopal Nelaturi, Ion Matei, Kumar Sricharan, “Classifying Heart Sound Recordings using Deep Convolutional Neural Networks and Mel-Frequency Cepstral Coefficients”, in *2016 Computing in Cardiology Conference (CinC)*, 813–816, 2016
- [7] Francois Chollet, “Building powerful image classification models using very little data” In *Keras*